

ARMY RESEARCH LABORATORY



The Effect of Appropriately and Inappropriately Applied Automation for the Control of Unmanned Systems on Operator Performance

**by Keryl Cosenzo, Raja Parasuraman, Krishna Pillalamarri,
and Theodric Feng**

ARL-TR-4933

September 2009

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

Army Research Laboratory

Aberdeen Proving Ground, MD 21005-5425

ARL-TR-4933

September 2009

The Effect of Appropriately and Inappropriately Applied Automation for the Control of Unmanned Systems on Operator Performance

Keryl Cosenzo, Krishna Pillalamarri, and Theodric Feng
Human Research and Engineering Directorate, ARL

Raja Parasuraman
George Mason University

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188
<p>Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p> <p>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</p>				
1. REPORT DATE (DD-MM-YYYY) September 2009	2. REPORT TYPE Final	3. DATES COVERED (From - To) October 2007–September 2008		
4. TITLE AND SUBTITLE The Effect of Appropriately and Inappropriately Applied Automation for the Control of Unmanned Systems on Operator Performance			5a. CONTRACT NUMBER	
			5b. GRANT NUMBER	
			5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Keryl Cosenzo, Raja Parasuraman,* Krishna Pillalamarri, and Theodric Feng			5d. PROJECT NUMBER 62716AH70	
			5e. TASK NUMBER	
			5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Laboratory ATTN: RDRL-HRS-E Aberdeen Proving Ground, MD 21005-5425			8. PERFORMING ORGANIZATION REPORT NUMBER ARL-TR-4933	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSOR/MONITOR'S ACRONYM(S)	
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.				
13. SUPPLEMENTARY NOTES *George Mason University, Fairfax, VA				
14. ABSTRACT Robotic technology will be a vital component of future combat. However, the combination of robotic operational tasks with other traditional military tasks will create high workload peaks during military operations. The objective of this research is to develop and evaluate flexible automation strategies to aid the operator in this complex military environment. In this experiment, we evaluated the effect of an automation that was invoked based on task load. Participants conducted a military reconnaissance mission using a simulation that required them to use an unmanned air vehicle sensor for target detection, monitor an unmanned ground vehicle, and respond to multi-level communications. Participants completed 16 missions in the environment, during which task load and automation were manipulated. The results of this experiment showed that operator performance did improve when the automation, an aided target-recognition system for the unmanned air vehicle, was invoked, relative to when it was not invoked. Further, when automation was appropriately applied (high task-load conditions), workload decreased significantly. This data, along with the results of other experiments discussed in this report, indicate that adaptive automation may be a useful mitigation strategy to help offset the potential deleterious effects of high cognitive load on U.S. Army robotic operators in a multitasking environment.				
15. SUBJECT TERMS design guidelines, robotics, simulation, unmanned systems, automation				
16. SECURITY CLASSIFICATION OF: a. REPORT Unclassified		17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 32	19a. NAME OF RESPONSIBLE PERSON Keryl Cosenzo
b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			19b. TELEPHONE NUMBER (Include area code) 410-278-5885

Contents

List of Figures	iv
List of Tables	iv
1. Introduction	1
2. Method	5
2.1 Participants	5
2.2 Instrumentation.....	5
2.3 Procedure.....	6
2.4 Experimental Design	6
3. Results	8
3.1 Objective Performance	8
3.1.1 UGV Route Planning Task Performance	8
3.1.2 Communications Task Performance	9
3.1.3 Change Detection Performance.....	9
3.1.4 Subjective Performance.....	11
3.2 Situation Awareness	12
4. Conclusions	13
5. References	15
Appendix A. Demographics Questionnaire	17
Appendix B. NASA-TLX	19
Appendix C. CC-SART	21
Distribution List	23

List of Figures

Figure 1. Robotic NCO simulation.....	3
Figure 2. Mean (standard error) reaction time for the UGV task.....	8
Figure 3. Mean (standard error) changes detected.....	10
Figure 4. Mean (standard error) workload scores.....	11
Figure 5. Mean (standard error) overall workload.....	12

List of Tables

Table 1. Counterbalancing scheme.....	7
Table 2. Mean (standard error) communication task performance in the low and high UAV task load condition.....	9
Table 3. Mean (standard error) change detection performance in the low and high UAV and communication task load conditions.....	10

1. Introduction

Robotic systems with diverse roles, tasks, and operating requirements are being designed for use in current and future military operations. Some of these systems will require active human control, whereas others may require high-level supervision with the possibility of intervention. No matter the level of autonomy on the robotic platform, Soldiers will be operating these systems while conducting other tasks that are essential to mission success, such as team communication and maintaining security. In the most extreme case, Soldiers will operate multiple heterogeneous robotics systems while on the move and while under enemy fire. In all cases, the workload will be variable—unpredictable—changing rapidly as a function of the military environment. Because of the likely increase in the cognitive workload demands on the Soldier, automation will be needed to support Soldier performance (Parasuraman et al., 2007). Automation refers to when a task that was previously completed by an operator is partially or entirely completed by a computer (Parasuraman et al., 2000). Levels of automation can range from the lowest level where the computer offers no assistance and the operator makes all the decisions to the highest level where the computer makes all the decisions without the human in the loop. To determine the appropriate level of automation, the effects on operator performance to include mental workload, situation awareness, and skill degradation need to be evaluated. Further, automation can have costs that need to be considered prior to implementation, such as loss of situation awareness and operator complacency. Research is needed in human-robot interaction to determine the optimum level of automation and to what functions or tasks it should be applied. The goal of this research is to investigate if and how automation should be implemented to support human supervision of unmanned systems so as to fully realize the benefits of force multiplication.

One way automation can be implemented is adaptively. In adaptive systems, the allocation of tasks between the human and machine agent is dynamic (Barnes et al., 2006). Adaptive automation uses mitigation criteria that drive an invocation mechanism to maintain an effective mixture of operator engagement and automation. The invocation mechanism is triggered by whatever measurement process is used to represent the current state of the operator and/or task. That measurement can be based on critical events, real-time operator performance, a predictive model of operator performance, or a hybrid method which combines one or more of these different invocation techniques (e.g., critical events and operator performance), so that the relative merits of each method can be maximized in order to minimize operator workload and maximize performance. The benefits of adaptive automation have been empirically demonstrated in several different domains to include basic multiple-task (i.e., the Multi-Attribute Task battery) and aviation environments (see Parasuraman [2000] for a review). The initial studies were designed to investigate whether the performance costs of certain forms of static automation, such as reduced situation awareness, complacency, skill degradation, etc., can be

mitigated by adaptive automation. Most of these studies used either a critical event or model-based approach to adaptive automation. A task was allocated dynamically to either human or machine control at some point in time during a simulated mission, either when some critical event occurred, or as dictated by a simple model of operator and system performance. For example, Hilburn et al. (1997) examined the effects of adaptive automation on the performance of military air traffic controllers who were provided with a decision aid for determining optimal descent trajectories of aircraft—a Descent Advisor (DA). The DA was either present at all times (static automation) or came on only when the traffic density exceeded a threshold. Hilburn et al. found significant benefits for controller workload (as assessed using pupillometric and heart rate variability measures) when the DA was provided adaptively during high traffic loads, compared to when it was available throughout (static automation) or only at low traffic loads. In addition to physiological measures of workload, other measures can also be used to assess the workload-leveling effect of adaptive automation. Kaber and Riley (1999), for example, used a secondary-task measurement technique to assess operator workload in a target acquisition task. They found that adaptive computer aiding based on the secondary-task measure enhanced performance on the primary task.

Only a small body of work has investigated the benefits (or potential costs) of flexible automation strategies, such as adaptive automation, in the context of robotic systems. There is limited work for the specific case of military unmanned systems and very little is known of the particular invocation methods that are best suited to this application case. The literature on automation has shown that the utilization and impact of automation on performance is not consistent across individuals (Lee and See, 2004). High workload, trust, and reliability are major factors in the efficacy of automation. When workload is high the operator may trade off situation awareness for decreasing his/her workload by depending on the automation aids, even in cases when it was not beneficial to do so. This is not universally true; in some cases automation improved overall performance even when the automated task required intervention because the operator's residual cognitive capacity was allocated effectively among the set of tasks (Galster and Parasuraman, 2003; Lorenz et al., 2002). Too often, the loss of situation awareness related to inefficient automation monitoring leads not only to performance decrements but also to an increasingly impoverished understanding of the work environment which can result in catastrophic errors over time (Endsley, 1996; Mosier and Skitka, 1996; Parasuraman and Riley, 1997). During complex multitasking situations there may be an increase in the trust and reliance on automation. Because of the uncertainty and risk associated with military environments, the use of automation must be done with extreme caution. If automation is inappropriately applied this can lead to performance decrements. For example, providing automation to an operator during a low task load period can cause complacency. However, returning a task to an operator (i.e., no automation) during this low task load phases can improve performance (Parasuraman et al., 1996). The Soldier and his/her chain of command need to maintain Situation Awareness (SA); keeping the Soldiers out of the loop will not only have

consequences for the immediate task but also predispose them to miss important cues signaling change (Endsley, 1996). Conversely, requiring Soldiers to engage in multiple tasks could very well have the same consequences.

We developed a program of research to examine the potential benefits of automation, specifically adaptive automation, on operator performance during complex unmanned systems operations. The first step was to develop the capability for evaluating human operator performance in managing unmanned air vehicles (UAVs) and unmanned ground vehicles (UGVs) so that subsequent evaluations of the efficacy of adaptive automation could be conducted. Cosenzo et al. (2006) developed the Robotic NCO simulation (figure 1) for these performance evaluations and subsequent automation experiments. The simulation requires operators to complete three military tasks from the same display space: UAV sensor use for target detection, UGV monitoring, and multi-level communications.

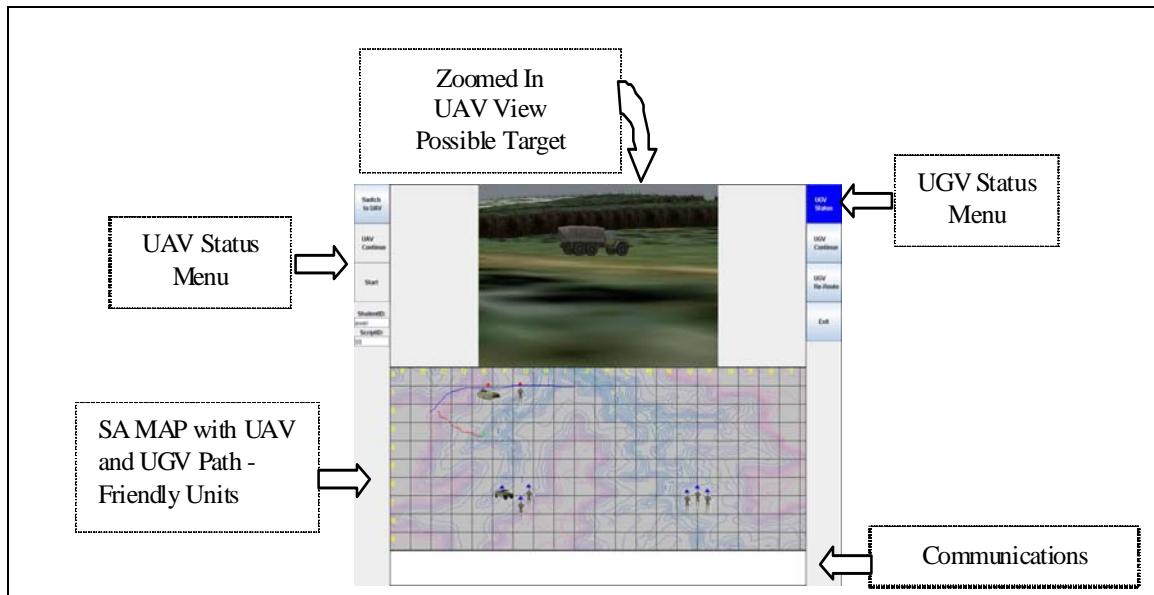


Figure 1. Robotic NCO simulation.

Cosenzo et al. (2006) used the Robotic NCO simulation environment to examine the multitasking requirements imposed on the operator and to identify tasks that were challenging and could benefit from automation. This experiment was used to establish a baseline level of unmanned system and multitask performance. Cosenzo et al. (2006) found that participants were good at integrating information received from the UAV and UGV during a simulated reconnaissance mission. However, the multi-tasking requirements of the Robotic NCO simulation decreased performance of the individual tasks that participants had to perform in addition to supervising the UVs. For example, communications task performance (e.g., reaction time) was degraded when task load was high (i.e., many UAV targets to process and many UGV stops to handle). Based on the results from the first experiment, Parasuraman et al. (2009) implemented an ancillary task, change detection embedded into the situation map of the Robotic

NCO display, as a potential trigger for adaptive automation. The change detection procedure required participants to identify when an icon changed location on the SA map. The rationale behind this procedure was that the transient and dynamic changes in the operator's SA could be captured by probing their awareness of the simulated battlefield environment via change detection performance. People often fail to notice changes in visual displays when they occur at the same time as various forms of visual transients (Simons and Ambinder, 2005; Simons and Rensink, 2005; Durlach, 2004). This change blindness phenomenon has been demonstrated in simple laboratory tasks and complex military tasks. In real life, a corollary exists in that in many tactical military operations the operator's situation display may often be updated without warning, so that the operator may miss the change. Thus, for the Robotic NCO experiments the hypothesis was that if automation could improve SA, then change detection performance should be enhanced with appropriately applied automation.

Using the change detection paradigm, Parasuraman et al. (2009) examined the effect of model-based and operator performance based automation on operator performance in the Robotic NCO simulation. The automation was an Aided Target Recognition System (ATR) that off-loaded the responsibility of identifying targets in the UAV imagery from the participants. The ATR was triggered by two methods, real-time assessment of operator change detection performance and a predictive model of operator performance. Results showed that compared to manual performance, both performance based and model based automation led to an increase in change detection accuracy and situation awareness and a decrease in workload. In comparison to the model based adaptive approach, there was a further increase in change detection accuracy and a concomitant reduction in workload with the performance based approach.

This current experiment complemented the efforts of Parasuraman et al. (2009) and examined the effect of task based automation on operator performance, change detection, in the Robotic NCO simulation. Automation was triggered in this experiment based on the level of task difficulty. More specifically, we examined the impact of applying an ATR when task load was high vs. when task load was low. We hypothesized that operator performance would be improved when the ATR was invoked relative to the conditions where the ATR was not invoked. Further, the ATR would be more effective when task load was high relative to when task load was low.

Hypotheses

Performance (i.e., reaction time and accuracy) will improve on the UGV and communications task when the ATR is invoked relative to the conditions where the ATR is not invoked.

The ATR will be more effective (i.e., increased task performance) when task load is high than when task load is low, that is when it is appropriately applied relative to inappropriately applied.

Workload and SA may improve when the ATR is invoked relative to the conditions where the ATR is not invoked.

2. Method

2.1 Participants

Twelve civilians (7 men and 5 women) with no prior experience with the Robotic NCO simulation participated in this study. The mean age was 32.6 years (range: 21–40). Participation in the experiment was strictly voluntary and participants could withdraw from the experiment at any time without penalty.

2.2 Instrumentation

The Robotic NCO simulation consisted of four military-relevant tasks:

1. UAV target identification task: A UAV flew following a series of pre-planned waypoints during the mission and received electronic hits from potential targets in the area, displayed as small white squares on the imagery. Participants had to locate and identify targets from UAV video imagery.
2. UGV route planning task: A UGV moved through the area following a series of pre-planned waypoints and during the mission the UGV stopped at obstacles seven times and requested help from the operator. Participants had to determine the appropriate course of action for the UGV, continue on pre-planned path or re-route the UGV around the obstacle.
3. Communications task with an embedded verbal SA probe task: Participants were presented messages both auditorily and visually. They had to monitor the messages and acknowledge the message when they heard their call sign. The messages also requested updates on the UGV/UAV status and the location of particular targets to assess SA.
4. Change detection task embedded within a situation map. At unpredictable times during the mission and after the situation map had been populated to a degree, an icon on the situation map (a target previously identified by the participants) changed its location. Participants pressed the space bar when they noticed the change.

The Demographics Questionnaire (appendix A) is a 10-item questionnaire that requests information regarding age, vision and hearing, military service, and computer experience.

NASA-Task Load Index (TLX) (appendix b): A subjective rating was given at the end of each mission on the participants' perceived workload as measured by the NASA-TLX questionnaire (Hart and Staveland, 1987). The NASA-TLX is a multi-dimensional rating procedure that derives an overall workload score based on a weighted average of ratings on six subscales (Mental Demand, Physical Demand, Temporal Demand, Own Performance, Effort, and Frustration).

Cognitive Compatibility Situation Awareness Technique (CC-SART) Questionnaire (appendix C): A subjective rating of perceived SA was given at the end of each mission, as measured by the CC-SART Questionnaire (Taylor, 1990). Participants rated their experience with the task on three dimensions: Activation of Knowledge, Ease of Reasoning, and Level of Processing. Level of Processing is the degree to which the situation involves, at the low level, natural automatic, intuitive, and associated processing, or at the high level, analytic, considered, conceptual, and abstract processing. Ease of Reasoning is the degree to which the situation, at the low level, is confusing and contradictory, or, at the high level, is straightforward and understandable. Activation of Knowledge is the degree to which the situation, at the low level, is strange and unusual, or, at the high level is recognizable and familiar.

2.3 Procedure

Prior to the start of the experiment, the participants were briefed on the purpose and procedures of the study and read the Volunteer Agreement Affidavit. They then completed the Demographics Questionnaire. The Robotic NCO simulation was described to the participants and functionality of each button in the simulation was explained.

Participants were asked to take the role of a robotic operator in a Mounted Combat System company (MCS). The participant conducted a reconnaissance mission for the MCS platoon. The MCS platoon was allocated an Armed Reconnaissance Vehicle (ARV) and an information feed via an UAV. Therefore, to replicate actual battlefield assets, the participant used two robotic systems, a UGV and a UAV, to complete the mission. Intelligence has identified areas of interest (designated on the common operational picture map with letters) and the UGV had been waypoint-planned to look at these named areas of interest. The UAV had also been waypoint-planned to view around the areas.

Participants were told that the UAV, UGV, and verbal SA communications tasks in this simulation were coordinated tasks that supported the overall goal—a reconnaissance mission in which participants had to be aware of friendly and enemy unit movements and of the positions of their UAV and UGV assets. The participants completed two training missions to ensure that they understood the task. After training, the participants completed 16 missions. Following each mission they completed the NASA-TLX and the CC-SART.

2.4 Experimental Design

The experiment was a $2 \times 2 \times 2$ within subjects design. The within-subjects factors were the number of targets to be identified in the UAV Task (low and high), Communications Task (Low and High) and Levels of Automation (No Automation and Automation). For the High and Low UAV Task conditions, 20 and 10 targets were presented in the UAV view, respectively. For the Communication task, there were 25 communications received during a mission, 20 callsigns and five status questions. For the High and Low conditions, the participants were presented 16 callsigns to acknowledge (out of 20) and four callsigns to acknowledge (out of 20),

respectively. In each condition the remaining five communications events were status questions about the UGV, UAV, and enemy position. In the Automation condition, an ATR system was invoked at the beginning of the mission, and the participants were not responsible for identifying targets in the UAV view. Participants were instructed to monitor the results of the ATR, since they were required at the end of the mission to evaluate the best platoon path to take following the reconnaissance mission. In the No Automation condition, the ATR was not active and the participants manually identified UAV targets. The ATR was invoked in both the high and low task load conditions (i.e., UAV and communications) to assess performance during appropriately and inappropriately applied automation.

Participants completed 16 missions (two of each type). Each mission was 5 min in duration and the order of the missions was counterbalanced (see table 1).

Table 1. Counterbalancing scheme.

Participants	Order							
	Mission							
	1	2	3	4	5	6	7	8
1,2,3	High UAV Low Comms No Automation	Low UAV Low Comms Automation	High UAV High Comms No Automation	Low UAV High Comms Automation	High UAV Low Comms No Automation	Low UAV Low Comms Automation	High UAV High Comms No Automation	Low UAV High Comms Automation
4,5,6	High UAV High Comms No Automation	Low UAV Low Comms Automation	High UAV High Comms Automation	Low UAV Low Comms Automation	High UAV High Comms No Automation	Low UAV Low Comms No Automation	High UAV High Comms Automation	Low UAV Low Comms Automation
7,8,9	Low UAV Low Comms Automation	High UAV High Comms Automation	Low UAV Low Comms No Automation	High UAV High Comms No Automation	Low UAV Low Comms Automation	High UAV High Comms Automation	Low UAV Low Comms No Automation	High UAV High Comms No Automation
10,11,12	Low UAV High Comms Automation	High UAV High Comms No Automation	Low UAV Low Comms Automation	High UAV Low Comms No Automation	Low UAV High Comms Automation	High UAV High Comms No Automation	Low UAV Low Comms Automation	High UAV Low Comms No Automation
Order								
Participants	Mission							
	9	10	11	12	13	14	15	16
1,2,3	Low UAV Low Comms No Automation	High UAV Low Comms Automation	Low UAV High Comms No Automation	High UAV High Comms Automation	Low UAV Low Comms No Automation	High UAV Low Comms Automation	Low UAV High Comms No Automation	High UAV High Comms Automation
4,5,6	Low UAV High Comms No Automation	High UAV Low Comms No Automation	Low UAV High Comms Automation	High UAV Low Comms Automation	Low UAV High Comms No Automation	High UAV Low Comms No Automation	Low UAV High Comms Automation	High UAV Low Comms Automation
7,8,9	High UAV Low Comms Automation	Low UAV High Comms Automation	High UAV Low Comms No Automation	Low UAV High Comms No Automation	High UAV Low Comms Automation	Low UAV High Comms Automation	High UAV Low Comms No Automation	Low UAV High Comms No Automation
10,11,12	High UAV High Comms Automation	Low UAV High Comms No Automation	High UAV Low Comms Automation	Low UAV Low Comms No Automation	High UAV High Comms Automation	Low UAV High Comms No Automation	High UAV Low Comms Automation	Low UAV Low Comms No Automation

3. Results

Multivariate analyses of variance (MANOVA) were conducted to examine the effects of UAV Task Load, Communication Task Load, and Automation on performance, workload, and SA. Performance on the tasks was measured as reaction time and percent correct. Workload and SA were measured by the NASA-TLX and the CC-SART, respectively. Performance data for the UAV task was not analyzed since the participants did not have to identify targets in the UAV task for half of the missions.

3.1 Objective Performance

3.1.1 UGV Route Planning Task Performance

Figure 2 is a graph of the mean (standard error) reaction time for the UGV Task. Results showed that reaction times were shorter in the automation than the no-automation conditions. The decrease in reaction times was larger in the high UAV condition than the low UAV conditions.

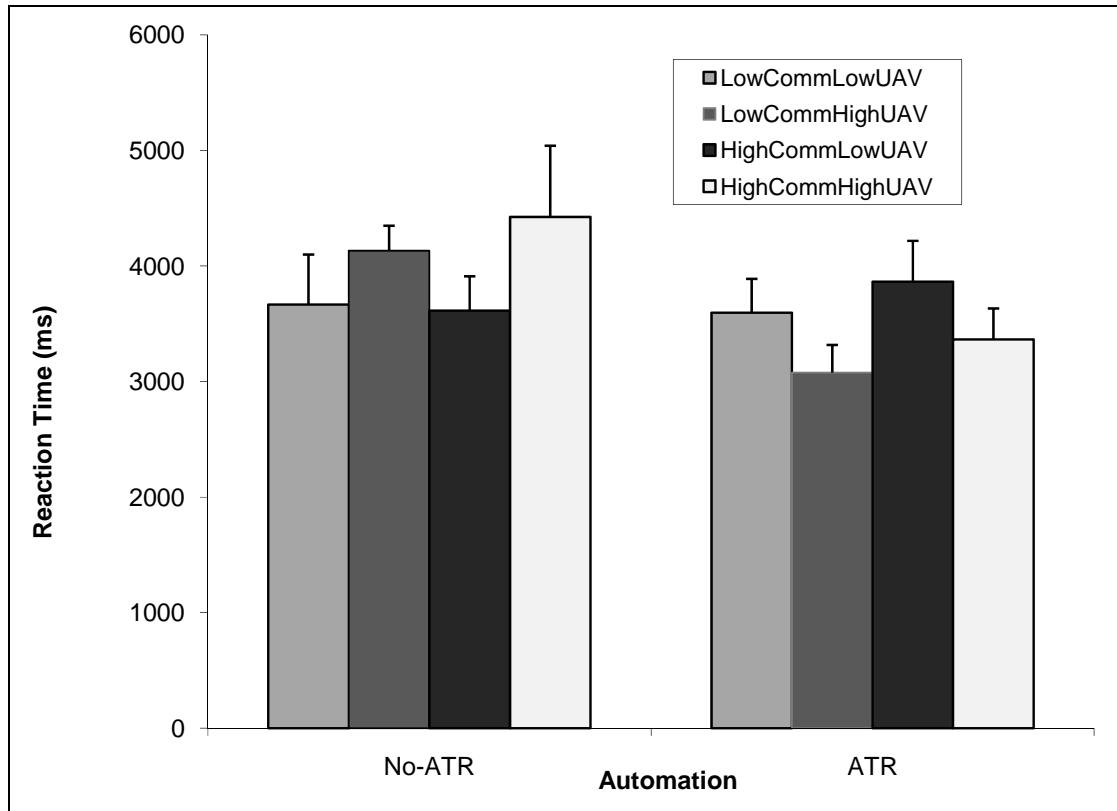


Figure 2. Mean (standard error) reaction time for the UGV task.

A MANOVA was conducted for UGV task performance. Results for UGV task performance showed a significant interaction of Level of Automation \times UAV Task \times Communication Task, $F(2,10) = 3.91, p < .05$. The three way interaction was significant for percent correct, $F(1,11) = 7.85, p < .01$ but not for reaction time, $p > .10$. There was also a significant interaction of Level of Automation \times UAV task, $F(2,10) = 6.49, p < .01$. The two way interaction was significant for reaction time, $F(1,11) = 9.51, p < .01$ but not for percent correct, $p > .10$.

To resolve the interaction of Level of Automation \times UAV Task \times Communication Task separate ANOVAs were conducted for each level of Automation, Communication Task, and UAV Task. All three ANOVAs yielded non-significant two-way interactions. Thus, the three-way interaction was not resolvable.

3.1.2 Communications Task Performance

Table 2 shows the means (standard errors) for communication task performance. Performance was poorer in the high UAV condition than the low UAV condition, that is reaction times were higher and accuracy was lower.

A MANOVA was conducted for Communications task performance. Results for Communications task performance showed a significant main effect of UAV task load, $F(2,10) = 12.78, p < .01$. The main effect for UAV task load was significant for percent correct and reaction time, $F(1,11) = 6.62, p < .02$ and $F(1,11) = 14.4, p < .01$, respectively. No other main effects or interactions were significant.

Table 2. Mean (standard error) communication task performance in the low and high UAV task load condition.

Variable	Level of UAV Task Load	Mean (SE)
Percent Correct	Low UAV	94.79 (1.76)
	High UAV	90.88 (1.76)
Reaction Time	Low UAV	2291.59 (107.35)
	High UAV	2439.08 (89.97)

3.1.3 Change Detection Performance

Figure 3 is a graph of the mean (standard error) for the percent of changes detected. An ANOVA was conducted for change detection performance. Results for change detection showed that more icon changes were detected in the Automation condition than the No Automation condition. This pattern was significant for the Low Communication condition but not the High condition. For change detection the interaction of Level of Automation \times Communication Task and the main effect for Communications Task were significant, $F(1, 11) = 5.80, p < .03$ and $F(1, 11) = 7.30, p < .02$. Subsequent ANOVAs showed a significant effect of Level of Automation for the Low Communication condition, $F(1, 11) = 5.88, p < .03$ but not the High condition, $p > .10$.

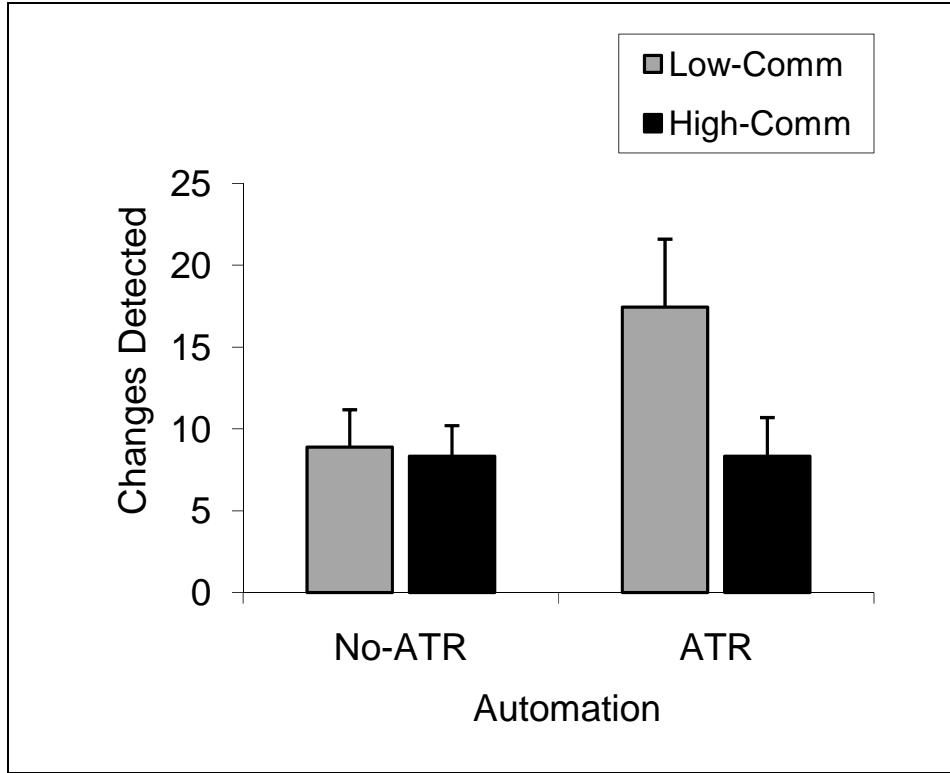


Figure 3. Mean (standard error) changes detected.

Results also showed that more icon changes were detected in the Low Communication than the High Communication condition when the number of UAV targets was high but not when the number of targets was low (see table 3). More specifically, for change detection, the interaction of UAV Task \times Communication Task was significant, $F(1, 11) = 10.3, p < .01$. Subsequent ANOVAs showed a significant effect of Communications Task for the High UAV Task condition, $F(1, 11) = 12.24, p < .01$ but not the Low Task condition, $p > .10$. No other interactions or main effects were significant.

Table 3. Mean (standard error) change detection performance in the low and high UAV and communication task load conditions.

Level of Communication Task Load	Level of UAV Task Load	Mean (SE)
Low	Low UAV	11.97 (2.76)
	High UAV	15.36 (3.53)
High	Low UAV	9.63 (2.50)
	High UAV	7.03 (1.92)

3.1.4 Subjective Performance

Workload: figures 4 and 5 are graphs of the mean (standard error) workload scores on the subscales and overall score of the NASA-TLX.

A MANOVA was conducted for the subscales of the NASA-TLX. Results for subjective workload showed a significant main effect for Level of Automation, $F(6,6) = 8.72, p < .01$. The main effect for Level of Automation was significant for mental demand, temporal demand, frustration, effort and performance, $F(1,11) = 41.1, p < .01$, $F(1,11) = 35.1, p < .01$, $F(1,11) = 4.74, p < .05$, $F(1,11) = 12.09, p < .01$., and $F(1,11) = 15.22, p < .01$, respectively.

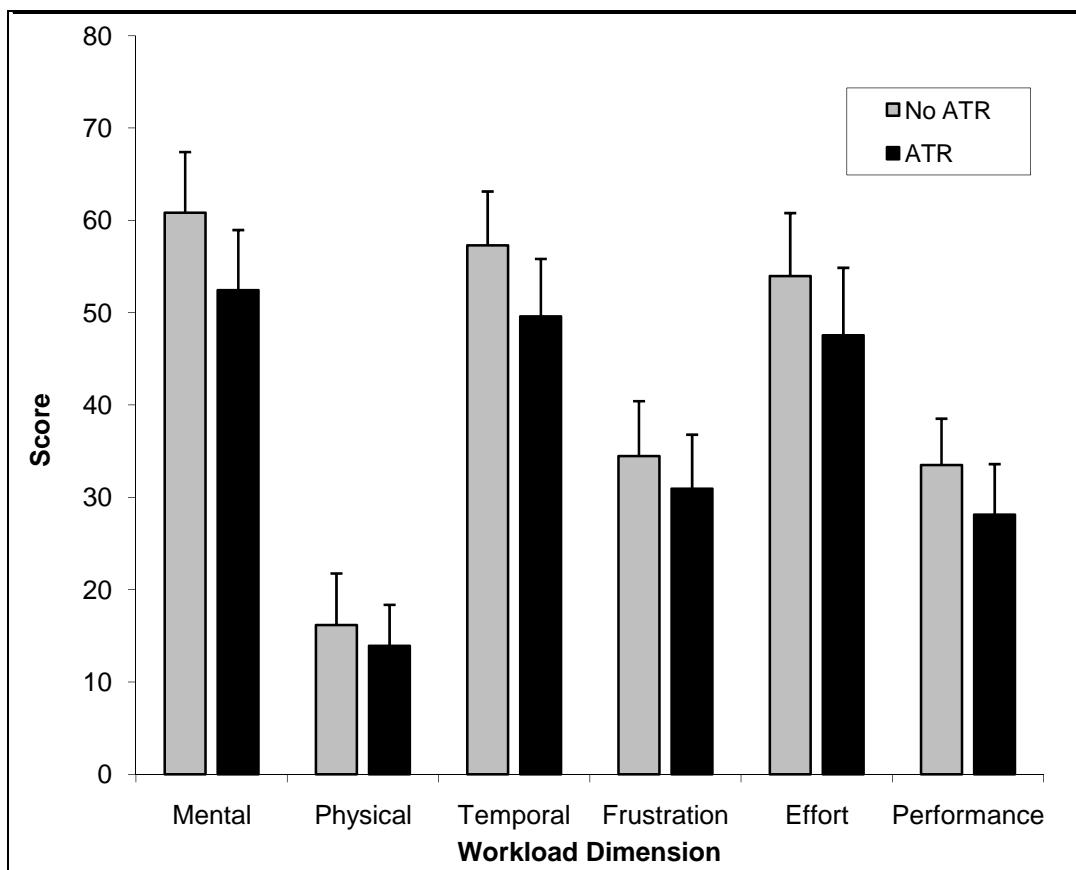


Figure 4. Mean (standard error) workload scores.

A second workload analysis was conducted for overall workload (i.e., composite NASA-TLX). Overall workload was lower in the Automation condition than the No Automation condition. Additionally, the decrease in workload between the Automation and No Automation condition was greater when the number of UAV targets was high relative to when the number of targets was low (see figure 5). More specifically, for overall workload the interaction of Level of

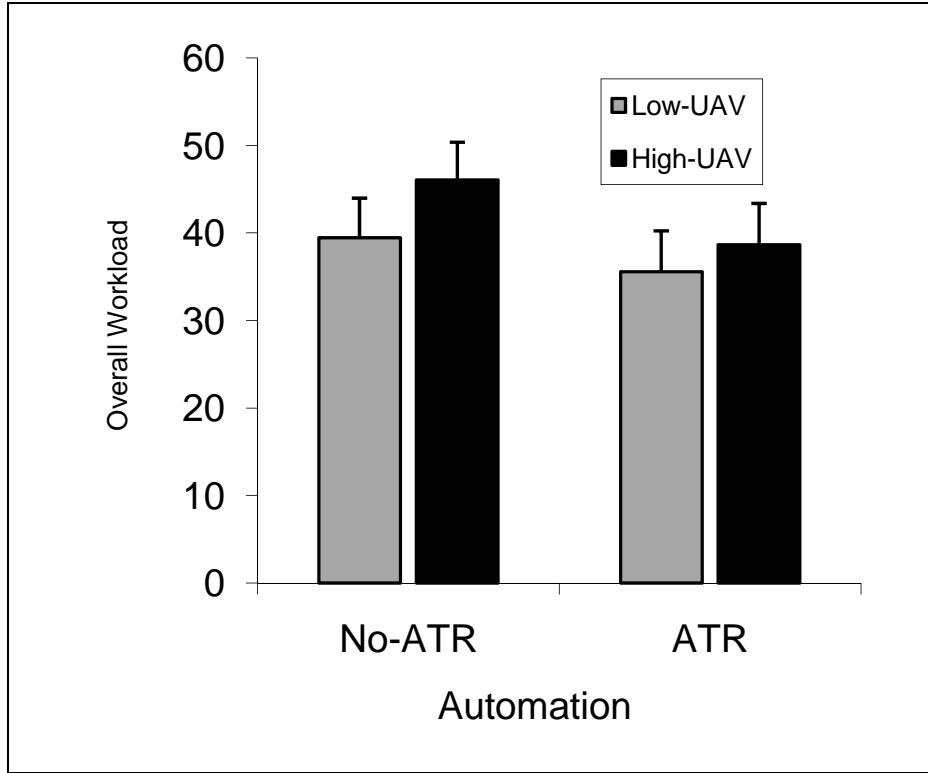


Figure 5. Mean (standard error) overall workload.

Automation \times UAV Task and the main effects of Automation and UAV Task were significant, $F(1, 11) = 12.85, p < .01$, $F(1, 11) = 27.0, p < .01$, and $F(1, 11) = 11.84, p < .01$, respectively.

Subsequent ANOVAs showed a significant effect of Automation for the High UAV condition and Low UAV Condition, $F(1, 11) = 27.5, p < .01$ and $F(1, 11) = 17.9, p < .01$.

No other main effects or interactions of UAV Task Load, Communication Task Load, and Automation were significant.

3.2 Situation Awareness

A MANOVA was conducted for the subscales of the CC-SART. Results for perceived SA, Level of Processing, showed that the interaction of Level of Automation \times UAV Task and the main effect of Automation were significant, $F(1, 10) = 6.14, p < .03$ and $F(1, 10) = 9.72, p < .01$. Subsequent ANOVAs showed a significant effect of Level of Automation for the High UAV condition, $F(1, 10) = 25.4, p = .01$, but not the Low UAV condition, $p > .10$. In the High UAV condition, Level of Processing was higher in the No Automation condition ($\bar{x} = 3.88, S.E. = .25$) than the Automation condition ($\bar{x} = 3.20, S.E. = .21$). There were no significant differences for Activation of Knowledge and Ease of Reasoning, $p > .10$. No other main effects or interactions of UAV Task Load, Communication Task Load, and Automation were significant.

4. Conclusions

The objective of this experiment was to examine the effect of task based automation on operator performance, change detection, in the Robotic NCO simulation. More specifically, we examined the impact of applying an ATR when task load was high vs. when task load was low. This experiment complemented the work of Parasuraman et al. (2009). Parasuraman and colleagues (2009) used the same research paradigm and automation (an ATR) which was triggered based on an individual's change detection performance. When performance dropped below a threshold, the ATR was triggered. In contrast, in this experiment the automation was triggered irrespective of individual performance. The ATR was either always on or off. We examined the benefit of automation under high task load and low task load conditions. The literature on automation suggests that the appropriate application of automation (i.e., only during peak task load) can enhance operation performance, SA, and a decrease workload. In appropriately applied automation can lead to unintended performance degradation.

In this experiment, we hypothesized that the ATR would be more effective (i.e., increased task performance) when task load is high than when task load is low, that is when it is appropriately applied relative to inappropriately applied. The results of this experiment partially supported our hypotheses. Operator performance (i.e., change detection) did improve when the ATR was invoked relative to when it was not invoked. In addition, ATR was more effective in the low communications task condition (high task load). This result, though seemingly contrary, is consistent with findings by Cosenzo et al. (2006) who showed that when the uncertainty of communications was high (low communications load) participants took longer to respond to communications when they had many UAV targets and UGV requests. The high-priority but infrequently occurring communications pose a particularly high monitoring load on the operator, and as a result this task condition benefits from automation. The automation enhanced the operator's situation awareness of the icon changes on his SA map during the high task load period. A similar result emerged for performance on the UGV task. The operator responded more quickly to the UGV when the ATR was invoked relative to when it was not. Further, there was a decrease in response time between the non-ATR and ATR conditions, when the task load of communication task was high (low communications load). Automation condition did not significantly impact communications task performance directly, although the loading of this task did affect the other aspects of the Robotic NCO environment.

Further, this experiment showed that when automation was appropriately applied (high task load conditions) workload and SA decreased significantly. Task load did not affect self-reported workload. This may be due to the fact that the Robotic NCO environment is challenging. The amount of effort expended to complete the low and high task load conditions are still high.

The magnitude of the effect of appropriately applied static automation is not as large as that reported by Cosenzo et al. (2006) and Parasuraman et al. (2009) for adaptive automation. Those studies, when compared to this one conducted with the Robotic NCO simulation, indicate that automation may be a useful mitigation strategy to help offset the potential deleterious effects of high cognitive load on U.S. Army robotic operators in a multitasking environment. The data also suggest an advantage for adaptively automating a task vs. statically automation. Through adaptive automation, we can take into consideration individual differences in performance and apply automation only on an as needed basis. When performance stabilizes, the automation can be revoked and the operator can re-engage in that task. As we further understand the efficacy of adaptable or adaptive options in multitasking environments, result from simulations should transitioned into increasingly realistic simulations.

5. References

Barnes, M.; Parasuraman, R.; Cosenzo, K. Adaptive Automation for Military Robotic Systems. *RTOTR-HFM-078 Uninhabited Military Vehicles: Human Factors Issues in Augmenting the Force*; NATO Technical Report; 2006; pp 420–440.

Cosenzo, K.; Parasuraman, R.; Barnes, M.; Novak, A. *Adaptive Automation for Robotic Military Systems*; ARL-TR-3808; U.S. Army Research Laboratory: Aberdeen Proving Ground, MD, 2006.

Durlach, P. Change Blindness and Its Implications for Complex Monitoring and Control Systems Design and Operator Training. *Human-Computer Interaction* **2004**, *19* (4), 423–451.

Endsley, M. Automation and Situation Awareness. *Automation and Human Performance: Theory and Application. Human Factors in Transportation*; Lawrence Erlbaum Associates: Mahwah, NJ, 1996; pp 163–179.

Galster, S. M.; Parasuraman, R. The Application of a Qualitative Model of Human – Interaction With Automation in a Complex Flight Task. In *Proceedings of the 12th International Symposium of Aviation Psychology*, Ohio State University: Dayton, OH, 2003; pp 411–416.

Hart, S. G.; Staveland, L. E. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*; Hancock, P. A., Meshkati, N., Eds.; Elsevier: Amsterdam, 1987.

Hilburn, B.; Jorna, P. G.; Byrne, E. A.; Parasuraman, R. The Effect of Adaptive Air Traffic Control (ATC) Decision Aiding on Controller Mental Workload. *Human-Automation Interaction: Research and Practice*; Mouloua, M., Koonce, J., Eds.; Erlbaum Associates: Mahwah, NJ, 1997; pp 84–91.

Kaber, D.; Riley, J. Adaptive Automation of a Dynamic Control Task Based on Workload Assessment Through a Secondary Monitoring Task. *Automation Technology and Human Performance: Current Research and Trends*; Scerbo, M., Mouloua, M., Eds.; Lawrence Erlbaum Associates: Mahwah, NJ, 1999; pp 129–133.

Lee, J. D.; See, K. A. Trust in Automation: Designing for Appropriate Reliance. *Human Factors* **2004**, *46* (1), 50–80.

Lorenz, B.; Di Nocera, F.; Rottger, S.; Parasuraman, R. Automated Fault Management in a Simulated Space Flight Micro-World. *Aviation, Space and Environmental Medicine* **2002**, *73*, 886–897.

Mosier K. L.; Skitka, L. J. Human Decision Makers and Automated Decision Aids: Made From Each Other? *Automation and Human Performance: Theory and Application. Human Factors in Transportation*; Parasuraman, R., Mouloua, M., Eds.; Lawrence Erlbaum Associates: Mahway, NJ, 1996; pp 163–176.

Parasuraman, R. Designing Automation for Human Use: Empirical Studies and Quantitative Models. *Ergonomics* **2000**, *43*, 931–951.

Parasuraman, R.; Riley, V. Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors* **1997**, *39*, 230–253.

Parasuraman, R.; Mouloua, M.; Molloy, R. Effects of Adaptive Task Allocation on Monitoring of Automated Systems. *Human Factors* **1996**, *38*, 665–679.

Parasuraman, R.; Sheridan, T. B.; Wickens, C. D. A Model of Types and Levels of Human Interaction With Automation. *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans* **2000**, *30*, 286–297.

Parasuraman, R.; Barnes, M.; Cosenzo, K. Adaptive Automation for Human Robt Teaming in Future Command and Control Systems. *International Journal of Command and Control* **2007**, *1* (2), 43–68.

Parasuraman, R.; Cosenzo, K. A.; de Visser, E. Adaptive Automation for Human Supervision of Multiple Uninhabited Vehicles: Effects on Change Detection, Situation Awareness, and Workload. *Military Psychology* **2009**, *21*, 270–297.

Simons, D. J.; Ambinder, M. S. Change Blindness: Theory and Consequences. *Current Directions in Psychological Science* **2005**, *14* (1), 44–48.

Simons, D. J.; Rensink, R. A. Change Blindness: Past, Present, and Future. *Trends in Cognitive Sciences* **2005**, *9* (1), 16–20.

Taylor, R. M. Situational Awareness Rating Technique (SART): The Development of a Tool for Aircraft Systems Design. In *Situational Awareness in Aerospace Operations*; Neuilly Sur Seine, France: NATO-AGARD, AGARD-CP-478, 1990.

Appendix A. Demographics Questionnaire

This appendix appears in its original form, without editorial change.

Demographics Questionnaire

Participant ID _____

1. AGE: _____

2. GENDER: ___Male ___ Female

3. Do you wear glasses? ___ Yes ___ No

4. Is your vision corrected to 20/20 with eyeglasses or contacts? ___Yes ___ No

5. Do you have an apparent hearing impairment? ___Yes ___No

6. Are you in the Army? ___Yes ___No

If yes, for how many years? ___Less than 5 years ___5-10 years ___ 11-15 years ___ 16-20 years ___ 20 years or more

7. What is your rank? _____ What is your MOS? _____

8. How often do you use a computer?

___Never ___Daily ___Weekly ___Monthly ___Once or twice a year

9. Do you use the computer to play games? ___Yes ___No

If yes, how often? ___Daily ___Weekly ___Monthly ___Once or twice a year

10. Do you play console games (e.g. Playstation2, etc)? ___Yes ___No

If yes, how often? ___Daily ___Weekly ___Monthly ___Once or twice a year

Appendix B. NASA-TLX

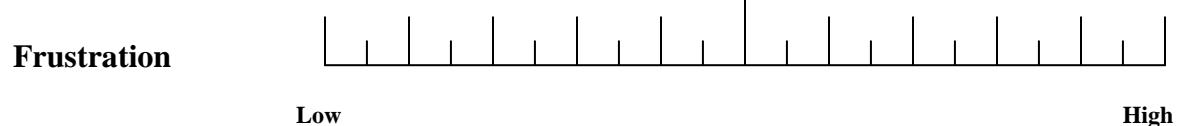
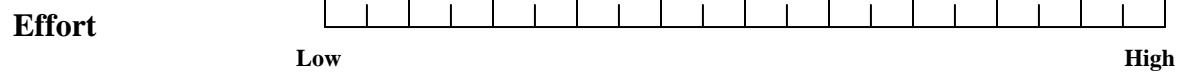
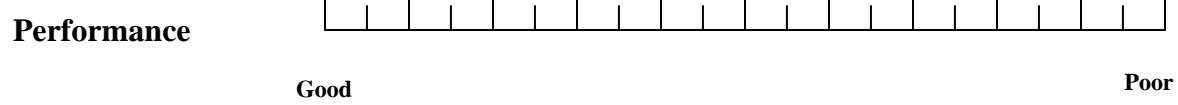
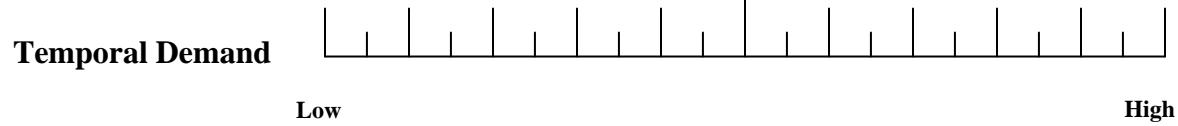
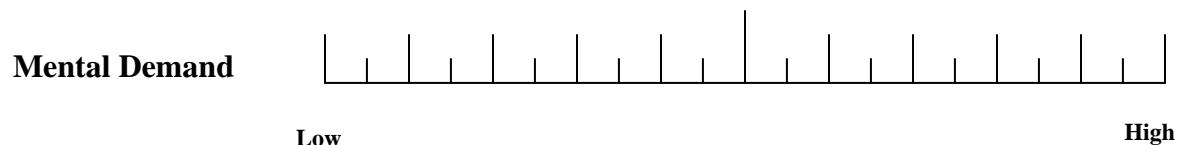
This appendix appears in its original form, without editorial change.

NASA TLX Questionnaire

Participant ID: _____

TLX Workload Scale

Please rate your workload by putting a mark on each of the six scales at the point which matches your experience.



Appendix C. CC-SART

This appendix appears in its original form, without editorial change.

CC-SART

Participant ID: _____

For each dimension below please place a mark under the rating value that matches your experience with the task you just completed.

Dimension	Rating							High
	Low	1	2	3	4	5	6	
<u>Level of Processing:</u> Degree to which the situation involves at the low level natural automatic, intuitive, and associated processing, or at the high level, analytic, considered, conceptual and abstract processing.								
<u>Ease of Reasoning:</u> Degree to which the situation at the low level, is confusing and contradictory, or, at the high level, is straightforward and understandable.								
<u>Activation of Knowledge:</u> Degree to which the situation at the low level, is strange and unusual, or , at the high level is recognizable and familiar.								

NO. OF
COPIES ORGANIZATION

1 DEFENSE TECHNICAL
(PDF INFORMATION CTR
only) DTIC OCA
8725 JOHN J KINGMAN RD
STE 0944
FORT BELVOIR VA 22060-6218

1 DIRECTOR
US ARMY RESEARCH LAB
IMNE ALC HRR
2800 POWDER MILL RD
ADELPHI MD 20783-1197

1 DIRECTOR
US ARMY RESEARCH LAB
RDRL CIM L
2800 POWDER MILL RD
ADELPHI MD 20783-1197

1 DIRECTOR
US ARMY RESEARCH LAB
RDRL CIM P
2800 POWDER MILL RD
ADELPHI MD 20783-1197

ABERDEEN PROVING GROUND

1 DIR USARL
RDRL CIM G (BLDG 4600)

<u>NO. OF COPIES</u>	<u>ORGANIZATION</u>	<u>NO. OF COPIES</u>	<u>ORGANIZATION</u>
1	ARMY RSCH LABORATORY – HRED RDRL HRM A J MARTIN MYER CENTER BLDG 2700 RM 2D311 FORT MONMOUTH NJ 07703-5601	1	ARMY RSCH LABORATORY – HRED RDRL HRM AP D UNGVARSKY POPE HALL BLDG 4709 BCBL 806 HARRISON DR FORT LEAVENWORTH KS 66027-2302
1	ARMY RSCH LABORATORY – HRED RDRL HRM C A DAVISON 320 MANSCEN LOOP STE 115 FORT LEONARD WOOD MO 65473	1	ARMY RSCH LABORATORY – HRED RDRL HRM AJ J HANSBERGER JFCOM FE 115 LAKEVIEW PKWY STE B SUFFOLK VA 23435
1	ARMY RSCH LABORATORY – HRED RDRL HRM DI T DAVIS BLDG 5400 RM C242 REDSTONE ARSENAL AL 35898-7290	1	ARMY RSCH LABORATORY – HRED RDRL HRM DQ M R FLETCHER NATICK SOLDIER CTR AMSRD NSC WS E BLDG 3 RM 343 NATICK MA 01760-5020
1	COMMANDANT USAADASCH ATSA CD RDRL HRM DE DR HAWLEY 5800 CARTER RD FORT BLISS TX 79916-3802	1	ARMY RSCH LABORATORY – HRED RDRL HRM AT J CHEN 12423 RESEARCH PKWY ORLANDO FL 32826
1	ARMY RSCH LABORATORY – HRED RDRL HRS EA DR V J RICE BLDG 4011 RM 217 1750 GREELEY RD FORT SAM HOUSTON TX 78234-5002	1	ARMY RSCH LABORATORY – HRED RDRL HRM AT C KORTENHAUS 12350 RESEARCH PKWY ORLANDO FL 32826
1	ARMY RSCH LABORATORY – HRED RDRL HRM DG R SPINE BLDG 333 PICATINNY ARSENAL NJ 07806-5000	1	ARMY RSCH LABORATORY – HRED RDRL HRM AS C MANASCO SIGNAL TOWERS BLDG 29808A RM 303 FORT GORDON GA 30905-5233
1	ARMY RSCH LABORATORY – HRED ARMC FIELD ELEMENT RDRL HRM CH C BURNS THIRD AVE BLDG 1467B RM 336 FORT KNOX KY 40121	1	ARMY RSCH LABORATORY – HRED RDRL HRM CU 6501 E 11 MILE RD MS 284 BLDG 200A 2ND FL RM 2104 WARREN MI 48397-5000
1	ARMY RSCH LABORATORY – HRED AWC FIELD ELEMENT RDRL HRM DJ D DURBIN BLDG 4506 (DCD) RM 107 FORT RUCKER AL 36362-5000	1	ARMY RSCH LABORATORY – HRED FIRES CTR OF EXCELLENCE FIELD ELEMENT RDRL HRM AF C HERNANDEZ 3040 AUSTIN RD RM 221 FORT SILL OK 73503-9043
1	ARMY RSCH LABORATORY – HRED RDRL HRM CK J REINHART 10125 KINGMAN RD FORT BELVOIR VA 22060-5828	1	ARMY RSCH LABORATORY – HRED RDRL HRM AV S MIDDLEBROOKS 91012 STATION AVE RM 348 FORT HOOD TX 76544-5073
1	ARMY RSCH LABORATORY – HRED RDRL HRM AY M BARNES 2520 HEALY AVE STE 1172 BLDG 51005 FORT HUACHUCA AZ 85613-7069		

NO. OF
COPIES ORGANIZATION

1 ARMY RSCH LABORATORY – HRED
RDRL HRM DW E REDDEN
BLDG 4 CL 60
FORT BENNING GA 31905-5400

1 ARMY RSCH LABORATORY – HRED
RDRL HRM CN R SPENCER
DCSF DI HF
HQ USASOC BLDG E2929
FORT BRAGG NC 28310-5000

1 ARMY G1
(CD DAPE MR B KNAPP
only) 300 ARMY PENTAGON RM 2C489
WASHINGTON DC 20310-0300

ABERDEEN PROVING GROUND

11 DIR USARL
RDRL CIM G
S FOPPIANO
RDRL HR
T LETOWSKI
RDRL HRM B
J LOCKETT
RDRL HRM D
D HARRAH
RDRL HRS
L ALLENDER
RDRL HRS D
B AMREIN
RDRL HRS E
K COSENZO (5 CPS)

INTENTIONALLY LEFT BLANK.